

G. Bugaisky

#4



RAW SEQUENCE LISTING

DATE: 06/27/2002

PATENT APPLICATION: US/09/668,650

TIME: 14:02:04

Input Set : N:\Crf3\RULE60\09668650.raw

Output Set: N:\CRF3\06272002\I668650.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:  
5 (i) APPLICANT: Bosch, Hendrick J.  
6 Stiekema, Willem J.  
8 (ii) TITLE OF INVENTION: Hybrid Toxin  
10 (iii) NUMBER OF SEQUENCES: 15  
12 (iv) CORRESPONDENCE ADDRESS:  
13 (A) ADDRESSEE: Novartis Corporation  
14 (B) STREET: 3054 Cornwallis Road  
15 (C) CITY: Research Triangle Park  
16 (D) STATE: NC  
17 (E) COUNTRY: USA  
18 (F) ZIP: 27709  
20 (v) COMPUTER READABLE FORM:  
21 (A) MEDIUM TYPE: Floppy disk  
22 (B) COMPUTER: IBM PC compatible  
23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
26 (vi) CURRENT APPLICATION DATA:  
C--> 27 (A) APPLICATION NUMBER: US/09/668,650  
C--> 28 (B) FILING DATE: 22-Sep-2000  
29 (C) CLASSIFICATION:  
31 (vii) PRIOR APPLICATION DATA:  
33 (A) APPLICATION NUMBER: US/09/001,982  
34 (B) FILING DATE:  
36 (A) APPLICATION NUMBER: US 08/602,737  
37 (B) FILING DATE: 21-FEB-1996  
39 (viii) ATTORNEY/AGENT INFORMATION:  
40 (A) NAME: Meigs, J. Timothy  
41 (B) REGISTRATION NUMBER: 38,241  
42 (C) REFERENCE/DOCKET NUMBER: 130-4080/PCT/CIP  
44 (ix) TELECOMMUNICATION INFORMATION:  
45 (A) TELEPHONE: 919-541-8587  
46 (B) TELEFAX: 919-541-8689  
49 (2) INFORMATION FOR SEQ ID NO: 1:  
51 (i) SEQUENCE CHARACTERISTICS:  
52 (A) LENGTH: 3567 base pairs  
53 (B) TYPE: nucleic acid  
54 (C) STRANDEDNESS: single  
55 (D) TOPOLOGY: unknown  
57 (ii) MOLECULE TYPE: cDNA  
59 (iii) HYPOTHETICAL: NO  
C--> 61 (iv) ANTI-SENSE: NO

ENTERED

## RAW SEQUENCE LISTING

DATE: 06/27/2002

PATENT APPLICATION: US/09/668,650

TIME: 14:02:04

Input Set : N:\Crif3\RULE60\09668650.raw

Output Set: N:\CRF3\06272002\I668650.raw

```

63      (vi) ORIGINAL SOURCE:
64          (A) ORGANISM: Bacillus thuringiensis
66      (ix) FEATURE:
67          (A) NAME/KEY: CDS
68          (B) LOCATION: 1..3567
71      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
73 ATG GAG GAA AAT AAT CAA AAT CAA TGC ATA CCT TAC AAT TGT TTA AGT      48
74 Met Glu Glu Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
75   1           5           10           15
77 AAT CCT GAA GAA GTA CTT TTG GAT GGA GAA CGG ATA TCA ACT GGT AAT      96
78 Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn
79           20           25           30
81 TCA TCA ATT GAT ATT TCT CTG TCA CTT GTT CAG TTT CTG GTA TCT AAC      144
82 Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser Asn
83           35           40           45
85 TTT GTA CCA GGG GGA GGA TTT TTA GTT GGA TTA ATA GAT TTT GTA TGG      192
86 Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
87   50           55           60
89 GGA ATA GTT GGC CCT TCT CAA TGG GAT GCA TTT CTA GTA CAA ATT GAA      240
90 Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
91   65           70           75           80
93 CAA TTA ATT AAT GAA AGA ATA GCT GAA TTT GCT AGG AAT GCT GCT ATT      288
94 Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala Ile
95           85           90           95
97 GCT AAT TTA GAA GGA TTA GGA AAC AAT TTC AAT ATA TAT GTG GAA GCA      336
98 Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
99           100          105          110
101 TTT AAA GAA TGG GAA GAA GAT CCT AAT AAT CCA GAA ACC AGG ACC AGA      384
102 Phe Lys Glu Trp Glu Glu Asp Pro Asn Asn Pro Glu Thr Arg Thr Arg
103           115          120          125
105 GTA ATT GAT CGC TTT CGT ATA CTT GAT GGG CTA CTT GAA AGG GAC ATT      432
106 Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
107           130          135          140
109 CCT TCG TTT CGA ATT TCT GGA TTT GAA GTA CCC CTT TTA TCC GTT TAT      480
110 Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
111 145           150          155          160
113 GCT CAA GCG GCC AAT CTG CAT CTA GCT ATA TTA AGA GAT TCT GTA ATT      528
114 Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile
115           165          170          175
117 TTT GGA GAA AGA TGG GGA TTG ACA ACG ATA AAT GTC AAT GAA AAC TAT      576
118 Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn Tyr
119           180          185          190
121 AAT AGA CTA ATT AGG CAT ATT GAT GAA TAT GCT GAT CAC TGT GCA AAT      624
122 Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn
123           195          200          205
125 ACG TAT AAT CGG GGA TTA AAT AAT TTA CCG AAA TCT ACG TAT CAA GAT      672
126 Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
127           210          215          220
129 TGG ATA ACA TAT AAT CGA TTA CGG AGA GAC TTA ACA TTG ACT GTA TTA      720

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,650

DATE: 06/27/2002

TIME: 14:02:04

Input Set : N:\Crf3\RULE60\09668650.raw

Output Set: N:\CRF3\06272002\I668650.raw

130	Trp	Ile	Thr	Tyr	Asn	Arg	Leu	Arg	Arg	Asp	Leu	Thr	Leu	Thr	Val	Leu	
131	225					230					235					240	
133	GAT	ATC	GCC	GCT	TTC	CCA	AAC	TAT	GAC	AAT	AGG	AGA	TAT	CCA	ATT		768
134	Asp	Ile	Ala	Ala	Phe	Pro	Asn	Tyr	Asp	Asn	Arg	Arg	Tyr	Pro	Ile		
135					245				250						255		
137	CAG	CCA	GTT	GGT	CAA	CTA	ACA	AGG	GAA	GTT	TAT	ACG	GAC	CCA	TTA	ATT	816
138	Gln	Pro	Val	Gly	Gln	Leu	Thr	Arg	Glu	Val	Tyr	Thr	Asp	Pro	Leu	Ile	
139				260					265					270			
141	AAT	TTT	AAT	CCA	CAG	TTA	CAG	TCT	GTA	GCT	CAA	TTA	CCT	ACT	TTT	AAC	864
142	Asn	Phe	Asn	Pro	Gln	Leu	Gln	Ser	Val	Ala	Gln	Leu	Pro	Thr	Phe	Asn	
143			275					280					285				
145	GTT	ATG	GAG	AGC	AGC	GCA	ATT	AGA	AAT	CCT	CAT	TTA	TTT	GAT	ATA	TTG	912
146	Val	Met	Glu	Ser	Ser	Ala	Ile	Arg	Asn	Pro	His	Leu	Phe	Asp	Ile	Leu	
147		290				295					300						
149	AAT	AAT	CTT	ACA	ATC	TTT	ACG	GAT	TGG	TTT	AGT	GTT	GGA	CGC	AAT	TTT	960
150	Asn	Asn	Leu	Thr	Ile	Phe	Thr	Asp	Trp	Phe	Ser	Val	Gly	Arg	Asn	Phe	
151	305				310				315						320		
153	TAT	TGG	GGA	GGA	CAT	CGA	GTA	ATA	TCT	AGC	CTT	ATA	GGA	GGT	GGT	AAC	1008
154	Tyr	Trp	Gly	Gly	His	Arg	Val	Ile	Ser	Ser	Leu	Ile	Gly	Gly	Gly	Asn	
155				325					330					335			
157	ATA	ACA	TCT	CCT	ATA	TAT	GGA	AGA	GAG	GCG	AAC	CAG	GAG	CCT	CCA	AGA	1056
158	Ile	Thr	Ser	Pro	Ile	Tyr	Gly	Arg	Glu	Ala	Asn	Gln	Glu	Pro	Pro	Arg	
159			340					345					350				
161	TCC	TTT	ACT	TTT	AAT	GGA	CCG	GTA	TTT	AGG	ACT	TTA	TCA	AAT	CCT	ACT	1104
162	Ser	Phe	Thr	Phe	Asn	Gly	Pro	Val	Phe	Arg	Thr	Leu	Ser	Asn	Pro	Thr	
163			355				360						365				
165	TTA	CGA	TTA	TTA	CAG	CAA	CCT	TGG	CCA	GCG	CCA	CCA	TTT	AAT	TTA	CGT	1152
166	Leu	Arg	Leu	Leu	Gln	Gln	Pro	Trp	Pro	Ala	Pro	Pro	Phe	Asn	Leu	Arg	
167		370				375					380						
169	GGT	GTT	GAA	GGA	GTA	GAA	TTT	TCT	ACA	CCT	ACA	AAT	AGC	TTT	ACG	TAT	1200
170	Gly	Val	Glu	Gly	Val	Glu	Phe	Ser	Thr	Pro	Thr	Asn	Ser	Phe	Thr	Tyr	
171	385				390				395						400		
173	CGA	GGA	AGA	GGT	ACG	GTT	GAT	TCT	TTA	ACT	GAA	TTA	CCG	CCT	GAG	GAT	1248
174	Arg	Gly	Arg	Gly	Thr	Val	Asp	Ser	Leu	Thr	Glu	Leu	Pro	Pro	Glu	Asp	
175				405					410					415			
177	AAT	AGT	GTG	CCA	CCT	CGC	GAA	GGA	TAT	AGT	CAT	CGT	TTA	TGT	CAT	GCA	1296
178	Asn	Ser	Val	Pro	Pro	Arg	Glu	Gly	Tyr	Ser	His	Arg	Leu	Cys	His	Ala	
179			420					425						430			
181	ACT	TTT	GTT	CAA	AGA	TCT	GGA	ACA	CCT	TTT	TTA	ACA	ACT	GGT	GTA	GTA	1344
182	Thr	Phe	Val	Gln	Arg	Ser	Gly	Thr	Pro	Phe	Leu	Thr	Thr	Gly	Val	Val	
183			435				440						445				
185	TTT	TCT	TGG	ACG	CAT	CGT	AGT	GCA	ACT	CTT	ACA	AAT	ACA	ATT	GAT	CCA	1392
186	Phe	Ser	Trp	Thr	His	Arg	Ser	Ala	Thr	Leu	Thr	Asn	Thr	Ile	Asp	Pro	
187		450				455					460						
189	GAG	AGA	ATT	AAT	CAA	ATA	CCT	TTA	GTG	AAA	GGA	TTT	AGA	GTT	TGG	GGG	1440
190	Glu	Arg	Ile	Asn	Gln	Ile	Pro	Leu	Val	Lys	Gly	Phe	Arg	Val	Trp	Gly	
191	465				470				475						480		
193	GGC	ACC	TCT	GTC	ATT	ACA	GGA	CCA	GGA	TTT	ACA	GGA	GGG	GAT	ATC	CTT	1488
194	Gly	Thr	Ser	Val	Ile	Thr	Gly	Pro	Gly	Phe	Thr	Gly	Gly	Asp	Ile	Leu	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,650

DATE: 06/27/2002

TIME: 14:02:04

Input Set : N:\Crf3\RULE60\09668650.raw

Output Set: N:\CRF3\06272002\I668650.raw

195		485		490		495		
197	CGA AGA AAT ACC TTT GGT GAT TTT GTA TCT CTA CAA GTC AAT ATT AAT							1536
198	Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile Asn							
199		500		505		510		
201	TCA CCA ATT ACC CAA AGA TAC CGT TTA AGA TTT CGT TAC GCT TCC AGT							1584
202	Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser Ser							
203		515		520		525		
205	AGG GAT GCA CGA GTT ATA GTA TTA ACA GGA GCG GCA TCC ACA GGA GTG							1632
206	Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly Val							
207		530		535		540		
209	GGA GGC CAA GTT AGT GTA AAT ATG CCT CTT CAG AAA ACT ATG GAA ATA							1680
210	Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu Ile							
211	545		550		555		560	
213	GGG GAG AAC TTA ACA TCT AGA ACA TTT AGA TAT ACC GAT TTT AGT AAT							1728
214	Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser Asn							
215		565		570		575		
217	CCT TTT TCA TTT AGA GCT AAT CCA GAT ATA ATT GGG ATA AGT GAA CAA							1776
218	Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu Gln							
219		580		585		590		
221	CCT CTA TTT GGT GCA GGT TCT ATT AGT AGC GGT GAA CTT TAT ATA GAT							1824
222	Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile Asp							
223		595		600		605		
225	AAA ATT GAA ATT ATT CTA GCA GAT GCA ACA TTT GAA GCA GAA TCT GAT							1872
226	Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu Ala Glu Ser Asp							
227		610		615		620		
229	TTA GAA AGA GCA CAA AAG GCG GTG AAT GCC CTG TTT ACT TCT TCC AAT							1920
230	Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ser Asn							
231	625		630		635		640	
233	CAA ATC GGG TTA AAA ACC GAT GTG ACG GAT TAT CAT ATT GAT CAA GTA							1968
234	Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val							
235		645		650		655		
237	TCC AAT TTA GTG GAT TGT TTA TCA GAT GAA TTT TGT CTG GAT GAA AAG							2016
238	Ser Asn Leu Val Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys							
239		660		665		670		
241	CGA GAA TTG TCC GAG AAA GTC AAA CAT GCG AAG CGA CTC AGT GAT GAG							2064
242	Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu							
243		675		680		685		
245	CGG AAT TTA CTT CAA GAT CCA AAC TTC AGA GGG ATC AAT AGA CAA CCA							2112
246	Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Pro							
247		690		695		700		
249	GAC CGT GGC TGG AGA GGA AGT ACA GAT ATT ACC ATC CAA GGA GGA GAT							2160
250	Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp							
251	705		710		715		720	
253	GAC GTA TTC AAA GAG AAT TAC GTC ACA CTA CCG GGT ACC GTT GAT GAG							2208
254	Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly Thr Val Asp Glu							
255		725		730		735		
257	TGC TAT CCA ACG TAT TTA TAT CAG AAA ATA GAT GAG TCG AAA TTA AAA							2256
258	Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys							
259		740		745		750		

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,650

DATE: 06/27/2002

TIME: 14:02:04

Input Set : N:\Crif3\RULE60\09668650.raw

Output Set: N:\CRF3\06272002\I668650.raw

261	GCT	TAT	ACC	CGT	TAT	GAA	TTA	AGA	GGG	TAT	ATC	GAA	GAT	AGT	CAA	GAC	2304
262	Ala	Tyr	Thr	Arg	Tyr	Glu	Leu	Arg	Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	
263			755					760					765				
265	TTA	GAA	ATC	TAT	TTG	ATC	CGT	TAC	AAT	GCA	AAA	CAC	GAA	ATA	GTA	AAT	2352
266	Leu	Glu	Ile	Tyr	Leu	Ile	Arg	Tyr	Asn	Ala	Lys	His	Glu	Ile	Val	Asn	
267		770					775					780					
269	GTG	CCA	GGC	ACG	GGT	TCC	TTA	TGG	CCG	CTT	TCA	GCC	CAA	AGT	CCA	ATC	2400
270	Val	Pro	Gly	Thr	Gly	Ser	Leu	Trp	Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	
271	785					790				795						800	
273	GGA	AAG	TGT	GGA	GAA	CCG	AAT	CGA	TGC	GCG	CCA	CAC	CTT	GAA	TGG	AAT	2448
274	Gly	Lys	Cys	Gly	Glu	Pro	Asn	Arg	Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	
275				805					810						815		
277	CCT	GAT	CTA	GAT	TGT	TCC	TGC	AGA	GAC	GGG	GAA	AAA	TGT	GCA	CAT	CAT	2496
278	Pro	Asp	Leu	Asp	Cys	Ser	Cys	Arg	Asp	Gly	Glu	Lys	Cys	Ala	His	His	
279			820					825					830				
281	TCC	CAT	CAT	TTC	ACC	TTG	GAT	ATT	GAT	GTT	GGA	TGT	ACA	GAC	TTA	AAT	2544
282	Ser	His	His	Phe	Thr	Leu	Asp	Ile	Asp	Val	Gly	Cys	Thr	Asp	Leu	Asn	
283			835					840					845				
285	GAG	GAC	TTA	GGT	GTA	TGG	GTG	ATA	TTC	AAG	ATT	AAG	ACG	CAA	GAT	GGC	2592
286	Glu	Asp	Leu	Gly	Val	Trp	Val	Ile	Phe	Lys	Ile	Lys	Thr	Gln	Asp	Gly	
287	850					855						860					
289	CAT	GCA	AGA	CTA	GGG	AAT	CTA	GAG	TTT	CTC	GAA	GAG	AAA	CCA	TTA	TTA	2640
290	His	Ala	Arg	Leu	Gly	Asn	Leu	Glu	Phe	Leu	Glu	Glu	Lys	Pro	Leu	Leu	
291	865					870					875					880	
293	GGG	GAA	GCA	CTA	GCT	CGT	GTG	AAA	AGA	GCG	GAG	AAG	AAG	TGG	AGA	GAC	2688
294	Gly	Glu	Ala	Leu	Ala	Arg	Val	Lys	Arg	Ala	Glu	Lys	Lys	Trp	Arg	Asp	
295				885					890						895		
297	AAA	CGA	GAG	AAA	CTG	CAG	TTG	GAA	ACA	AAT	ATT	GTT	TAT	AAA	GAG	GCA	2736
298	Lys	Arg	Glu	Lys	Leu	Gln	Leu	Glu	Thr	Asn	Ile	Val	Tyr	Lys	Glu	Ala	
299			900						905					910			
301	AAA	GAA	TCT	GTA	GAT	GCT	TTA	TTT	GTA	AAC	TCT	CAA	TAT	GAT	AGA	TTA	2784
302	Lys	Glu	Ser	Val	Asp	Ala	Leu	Phe	Val	Asn	Ser	Gln	Tyr	Asp	Arg	Leu	
303			915					920					925				
305	CAA	GTG	GAT	ACG	AAC	ATC	GCG	ATG	ATT	CAT	GCG	GCA	GAT	AAA	CGC	GTT	2832
306	Gln	Val	Asp	Thr	Asn	Ile	Ala	Met	Ile	His	Ala	Ala	Asp	Lys	Arg	Val	
307	930					935					940						
309	CAT	AGA	ATC	CGG	GAA	GCG	TAT	CTG	CCA	GAG	TTG	TCT	GTG	ATT	CCA	GGT	2880
310	His	Arg	Ile	Arg	Glu	Ala	Tyr	Leu	Pro	Glu	Leu	Ser	Val	Ile	Pro	Gly	
311	945					950					955					960	
313	GTC	AAT	GCG	GCC	ATT	TTC	GAA	GAA	TTA	GAG	GGA	CGT	ATT	TTT	ACA	GCG	2928
314	Val	Asn	Ala	Ala	Ile	Phe	Glu	Glu	Leu	Glu	Gly	Arg	Ile	Phe	Thr	Ala	
315				965						970					975		
317	TAT	TCC	TTA	TAT	GAT	GCG	AGA	AAT	GTC	ATT	AAA	AAT	GGC	GAT	TTC	AAT	2976
318	Tyr	Ser	Leu	Tyr	Asp	Ala	Arg	Asn	Val	Ile	Lys	Asn	Gly	Asp	Phe	Asn	
319			980					985					990				
321	AAT	GGC	TTA	TTA	TGC	TGG	AAC	GTG	AAA	GGT	CAT	GTA	GAT	GTA	GAA	GAG	3024
322	Asn	Gly	Leu	Leu	Cys	Trp	Asn	Val	Lys	Gly	His	Val	Asp	Val	Glu	Glu	
323			995					1000					1005				
325	CAA	AAC	AAC	CAC	CGT	TCG	GTC	CTT	GTT	ATC	CCA	GAA	TGG	GAG	GCA	GAA	3072

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/668,650

DATE: 06/27/2002

TIME: 14:02:05

Input Set : N:\Crf3\RULE60\09668650.raw

Output Set: N:\CRF3\06272002\I668650.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:61 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:622 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:1176 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:1737 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]